

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/062,142

DATE: 04/23/98
TIME: 14:53:18

INPUT SET: S25251.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

ENTERED

SEQUENCE LISTING

1
2
3 (1) General Information
4
5 (i) APPLICANT: Sheppard, Paul O.
6
7 (ii) TITLE OF THE INVENTION: SERINE PROTEASE POLYPEPTIDES
8 AND MATERIALS AND METHODS FOR MAKING THEM
9
10 (iii) NUMBER OF SEQUENCES: 16
11
12 (iv) CORRESPONDENCE ADDRESS:
13 (A) ADDRESSEE: ZymoGenetics, Inc.
14 (B) STREET: 1201 Eastlake Avenue East
15 (C) CITY: Seattle
16 (D) STATE: WA
17 (E) COUNTRY: USA
18 (F) ZIP: 98102
19
20 (v) COMPUTER READABLE FORM:
21 (A) MEDIUM TYPE: Diskette
22 (B) COMPUTER: IBM Compatible
23 (C) OPERATING SYSTEM: DOS
24 (D) SOFTWARE: FastSEQ for Windows Version 2.0
25
26 (vi) CURRENT APPLICATION DATA:
27 (A) APPLICATION NUMBER:
28 (B) FILING DATE:
29 (C) CLASSIFICATION:
30
31 (vii) PRIOR APPLICATION DATA:
32 (A) APPLICATION NUMBER:
33 (B) FILING DATE:
34
35
36
37 (viii) ATTORNEY/AGENT INFORMATION:
38 (A) NAME: Parker, Gary E
39 (B) REGISTRATION NUMBER: 31,648
40 (C) REFERENCE/DOCKET NUMBER: 97-16
41
42 (ix) TELECOMMUNICATION INFORMATION:
43 (A) TELEPHONE: 206-442-6673
44 (B) TELEFAX: 206-442-6678
45 (C) TELEX:
46

09062142-04798

RAW SEQUENCE LISTING PATENT APPLICATION US/09/062,142

DATE: 04/23/98
TIME: 14:53:20

INPUT SET: S25251.raw

47

48

(2) INFORMATION FOR SEQ ID NO:1:

49

50

(i) SEQUENCE CHARACTERISTICS:

51

(A) LENGTH: 1634 base pairs

52

(B) TYPE: nucleic acid

53

(C) STRANDEDNESS: double

54

(D) TOPOLOGY: linear

55

56

(ix) FEATURE:

57

58

(A) NAME/KEY: Coding Sequence

59

(B) LOCATION: 105...1280

60

(D) OTHER INFORMATION:

61

62

(A) NAME/KEY: Signal Sequence

63

(B) LOCATION: 105...161

64

(D) OTHER INFORMATION:

65

66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

67

68

GGCACGAGGG GGAGCCGCGC GCTCTCTCCC GGCGCCCACA CCTGTCTGAG CGGCGCAGCG 60

69

AGCCGCGGCC CGGGCGGGCT GCTCGGCGCG GAACAGTGCT CGGC ATG GCA GGG ATT 116

70

Met Ala Gly Ile

71

72

73

CCA GGG CTC CTC TTC CTT CTC TTC TTT CTG CTC TGT GCT GTT GGG CAA 164

74

Pro Gly Leu Leu Phe Leu Leu Phe Phe Leu Leu Cys Ala Val Gly Gln

75

-15 -10 -5 1

76

77

GTG AGC CCT TAC AGT GCC CCC TGG AAA CCC ACT TGG CCT GCA TAC CGC 212

78

Val Ser Pro Tyr Ser Ala Pro Trp Lys Pro Thr Trp Pro Ala Tyr Arg

79

5 10 15

80

81

CTC CCT GTC GTC TTG CCC CAG TCT ACC CTC AAT TTA GCC AAG CCA GAC 260

82

Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn Leu Ala Lys Pro Asp

83

20 25 30

84

85

TTT GGA GCC GAA GCC AAA TTA GAA GTA TCT TCT TCA TGT GGA CCC CAG 308

86

Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser Ser Cys Gly Pro Gln

87

35 40 45

88

89

TGT CAT AAG GGA ACT CCA CTG CCC ACT TAC AAA GAA GCC AAG CAA TAT 356

90

Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Lys Glu Ala Lys Gln Tyr

91

50 55 60 65

92

93

CTG TCT TAT GAA ACG CTC TAT GCC AAT GGC AGC CGC ACA GAG ACN CAG 404

94

Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg Thr Glu Xaa Gln

95

70 75 80

96

97

GTG GGC ATC TAC ATC CTC AGC AGT AGT GGA GAT GGG GCC CAN CNC CGA 452

98

Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly Ala Xaa Xaa Arg

99

85 90 95

364740-21125050

RAW SEQUENCE LISTING PATENT APPLICATION US/09/062,142

DATE: 04/23/98
TIME: 14:53:22

INPUT SET: S25251.raw

100																		
101	GAC	TCA	GGG	TCT	TCA	GGA	AAG	TCT	CGA	AGG	AAG	CGG	CAG	ATT	TAT	GGC		500
102	Asp	Ser	Gly	Ser	Ser	Gly	Lys	Ser	Arg	Arg	Lys	Arg	Gln	Ile	Tyr	Gly		
103			100					105					110					
104																		
105	TAT	GAC	AGC	AGG	TTC	AGC	ATT	TTT	GGG	AAG	GAC	TTC	CTG	CTC	AAC	TAC		548
106	Tyr	Asp	Ser	Arg	Phe	Ser	Ile	Phe	Gly	Lys	Asp	Phe	Leu	Leu	Asn	Tyr		
107		115					120					125						
108																		
109	CCT	TTC	TCA	ACA	TCA	GTG	AAG	TTA	TCC	ACG	GGC	TGC	ACC	GGC	ACC	CTG		596
110	Pro	Phe	Ser	Thr	Ser	Val	Lys	Leu	Ser	Thr	Gly	Cys	Thr	Gly	Thr	Leu		
111		130					135					140				145		
112																		
113	GTG	GCA	GAA	AAN	CAT	GTC	CTC	ACA	GCT	GCC	CAC	TGC	ATA	CAC	GAT	GGA		644
114	Val	Ala	Glu	Xaa	His	Val	Leu	Thr	Ala	Ala	His	Cys	Ile	His	Asp	Gly		
115					150					155					160			
116																		
117	AAA	ACC	TAT	GTG	AAA	GGA	ACC	CAG	AAG	CTT	CGA	GTC	GGC	TTC	CTA	AAG		692
118	Lys	Thr	Tyr	Val	Lys	Gly	Thr	Gln	Lys	Leu	Arg	Val	Gly	Phe	Leu	Lys		
119				165					170					175				
120																		
121	CCC	AAG	TTT	AAA	GAT	GGT	GGT	CGA	GGG	GCC	AAC	GAC	TCC	ACT	TCA	GCC		740
122	Pro	Lys	Phe	Lys	Asp	Gly	Gly	Arg	Gly	Ala	Asn	Asp	Ser	Thr	Ser	Ala		
123			180					185					190					
124																		
125	ATG	CCC	GAG	CAG	ATG	AAA	TTT	CAG	TGG	ATC	CGG	GTG	AAA	CGC	ACC	CAT		788
126	Met	Pro	Glu	Gln	Met	Lys	Phe	Gln	Trp	Ile	Arg	Val	Lys	Arg	Thr	His		
127		195					200					205						
128																		
129	GTG	CCC	AAG	GGT	TGG	ATC	AAG	GGC	AAT	GCC	AAT	GAC	ATC	GGC	ATG	GAT		836
130	Val	Pro	Lys	Gly	Trp	Ile	Lys	Gly	Asn	Ala	Asn	Asp	Ile	Gly	Met	Asp		
131		210					215					220				225		
132																		
133	TAT	GAT	TAT	GCC	CTC	CTG	GAA	CTC	AAA	AAG	CCC	CAC	AAG	AGA	AAA	TTT		884
134	Tyr	Asp	Tyr	Ala	Leu	Leu	Glu	Leu	Lys	Lys	Pro	His	Lys	Arg	Lys	Phe		
135					230					235					240			
136																		
137																		
138	ATG	AAG	ATT	GGG	GTG	AGC	CCT	CCT	GCT	AAG	CAG	CTG	CCA	GGG	GGC	AGA		932
139	Met	Lys	Ile	Gly	Val	Ser	Pro	Pro	Ala	Lys	Gln	Leu	Pro	Gly	Gly	Arg		
140				245					250					255				
141																		
142	ATT	CAC	TTC	TCT	GGT	TAT	GAC	AAT	GAC	CGA	CCA	GGC	AAT	TTG	GTG	TAT		980
143	Ile	His	Phe	Ser	Gly	Tyr	Asp	Asn	Asp	Arg	Pro	Gly	Asn	Leu	Val	Tyr		
144			260					265					270					
145																		
146	CGC	TTC	TGT	GAC	GTC	AAA	GAC	GAG	ACC	TAT	GAC	TTG	TTG	TAC	CAG	CAA		1028
147	Arg	Phe	Cys	Asp	Val	Lys	Asp	Glu	Thr	Tyr	Asp	Leu	Leu	Tyr	Gln	Gln		
148		275					280					285						
149																		
150	TGC	GAT	GCC	CAG	CCA	GGG	GCC	AGC	GGG	TAT	GGG	GTA	TAT	GTG	AGG	ATG		1076
151	Cys	Asp	Ala	Gln	Pro	Gly	Ala	Ser	Gly	Tyr	Gly	Val	Tyr	Val	Arg	Met		
152		290					295				300					305		

Sequence 25251

INPUT SET: S25251.raw

[illegible]

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 392 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence
(B) LOCATION: 1...19
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

199	Met	Ala	Gly	Ile	Pro	Gly	Leu	Leu	Phe	Leu	Leu	Phe	Phe	Leu	Leu	Cys
200					-15					-10						-5
201	Ala	Val	Gly	Gln	Val	Ser	Pro	Tyr	Ser	Ala	Pro	Trp	Lys	Pro	Thr	Trp
202				1				5					10			
203	Pro	Ala	Tyr	Arg	Leu	Pro	Val	Val	Leu	Pro	Gln	Ser	Thr	Leu	Asn	Leu
204		15					20					25				
205	Ala	Lys	Pro	Asp	Phe	Gly	Ala	Glu	Ala	Lys	Leu	Glu	Val	Ser	Ser	Ser

[illegible]

RAW SEQUENCE LISTING PATENT APPLICATION US/09/062,142

DATE: 04/23/98
TIME: 14:53:25

INPUT SET: S25251.raw

```

206      30              35              40              45
207 Cys Gly Pro Gln Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Lys Glu
208              50              55              60
209 Ala Lys Gln Tyr Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg
210              65              70              75
211 Thr Glu Xaa Gln Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly
212              80              85              90
213 Ala Xaa Xaa Arg Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg
214              95              100              105
215 Gln Ile Tyr Gly Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe
216      110              115              120              125
217 Leu Leu Asn Tyr Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys
218              130              135              140
219 Thr Gly Thr Leu Val Ala Glu Xaa His Val Leu Thr Ala Ala His Cys
220              145              150              155
221 Ile His Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val
222              160              165              170
223 Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp
224              175              180              185
225 Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val
226      190              195              200              205
227 Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp
228              210              215              220
229
230 Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His
231              225              230              235
232 Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu
233              240              245              250
234 Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly
235              255              260              265
236 Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu
237      270              275              280              285
238 Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Tyr Gly Val
239              290              295              300
240 Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile
241              305              310              315
242 Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro
243              320              325              330
244 Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln
245              335              340              345
246 Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly Asp
247      350              355              360              365
248 Thr Val Phe Leu Pro Gly Ser Asn
249              370
250

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

Seq ID: 25251

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/062,142

DATE: 04/23/98
TIME: 14:53:27

INPUT SET: S25251.raw

Line

Error

Original Text

04/23/98 14:53:27